

Rich-club network topology to minimize synchronization cost due to phase difference among frequency-synchronized oscillators

Takamitsu Watanabe ^{*1}

¹*Department of Physiology, School of Medicine, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

Abstract

Functions of some networks, such as power grids and large-scale brain networks, rely on not only frequency synchronization, but also phase synchronization. Nevertheless, even after the oscillators reach to frequency-synchronized status, phase difference among oscillators often shows non-zero constant values. Such phase difference potentially results in inefficient transfer of power or information among oscillators, and avoid proper and efficient functioning of the network. In the present study, we newly define synchronization cost by the phase difference among the frequency-synchronized oscillators, and investigate the optimal network structure with the minimum synchronization cost through rewiring-based optimization. By using the Kuramoto model, we demonstrate that the cost is minimized in a network topology with rich-club organization, which comprises the densely-connected center nodes and peripheral nodes connecting with the center module. We also show that the network topology is characterized by its bimodal degree distribution, which is quantified by Wolfson's polarization index. Furthermore, we provide analytical interpretation on why the rich-club network topology is related to the small amount of synchronization cost.

* takawatanabe-tky@umin.ac.jp

I. INTRODUCTION

As power grids [1, 2] and networks of bursting neurons [3, 4], functions of some complex networks of oscillators are based on not only synchronization of frequencies of the oscillators, but also synchronization of their phases. However, in general, frequency synchronization is more achievable than phase synchronization. Phase difference among frequency-synchronized oscillators often falls into a non-zero constant, and such non-zero phase difference would avoid proper and efficient functioning of the complex networks.

In power grids, alternating voltage of the power plants in the grids should be synchronized around certain specific frequency (e.g., 50 Hz in the most parts of Europe and 60 Hz in the north America) [1], and disruption of the frequency synchronization causes a blackout in a large area [1, 5]. In addition, the phases of the voltages of the power plants are also required to be synchronized. As discussed in Appendix A, the difference in voltage phases among power plants inevitably causes power loss consumed as heat in power lines [1, 2]. In this sense, the phase difference in power grids can be regarded as synchronization cost. Considering recent increasing environmental awareness and soaring global demand of natural resources [6, 7], it is necessary to reduce the synchronization cost due to the phase difference among frequency-synchronized power plants.

Synchronization in large-scale brain networks also requires less difference in phase of neuronal activity among different brain regions. A series of prior experimental researches have shown that various important functions in large-scale brain networks are based on not only frequency- but also phase- synchronization [3, 4]. A previous electrophysiological study showed that spike activity recorded from monkeys' cortex exhibited phase synchronization in various frequency bands while the monkeys were conducting tasks that required integration of visual processing and motor responses [8]. In another careful electrophysiology study, Roelfsema and his colleagues recorded local field potentials (LFP) in the cerebral cortex in cats, and revealed that phase synchronization between LFPs recorded in the visual and parietal cortices was increased only when the cats focused their attention on stimuli [9]. Studies using electroencephalogram to record human brain activity found that increase of phase synchronization in various frequency was associated with learning and perception of images [10, 11]. Furthermore, the frequency- and phase- synchronization are considered to occur in large-scale brain networks [8, 12] with zero time-lag [4, 9, 13]. These researches suggest that phase synchronization in

large-scale brain networks is related to crucial functions such as integration of multiple information [3], neural communication [14], and spike-timing-dependent plasticity [4]. Actually, it is known that some types of the disruption of the synchronization cause dysfunctions of memory learning [15] or psychiatric disorders [16]. Considering these findings, it is to some extent reasonable to hypothesize that large-scale brain networks have a specific organization that minimizes phase differences among brain activity, and enables optimal phase synchronization in the entire networks.

These previous literatures indicate the importance to reduce phase difference among frequency-synchronized oscillators, which can be regarded as a type of cost that is spent during synchronization. However, little is examined about the optimal network topology that reduces this type of synchronization cost due to phase difference. Indeed, a series of previous literatures have investigated optimal network structures by introducing a different type of synchronization cost, which is needed for building or maintain of the optimal network infrastructure. A study regarded parameters based on coupling strength among oscillators as a cost, and demonstrated that homogeneous and uniform distribution of the coupling strength enhances the tendency of synchronization [17]. Another study employing the same definition of synchronization cost suggests that more heterogeneous network structures are required to simultaneously achieve both the maximum synchronizability and minimum synchronization cost based on coupling strength [18]. Another study revealed the optimal distribution of coupling strength that increases synchronizability among Kuramoto oscillators [19]. However, these prior researches have not focused on synchronization cost due to phase difference. Consequently, despite a line of prior researches on optimal conditions for synchrony in networks [20], it remains unclear about optimal network structures that minimize synchronization cost due to phase difference among oscillators.

In the present study, therefore, we examine the optimal network topology to minimize the phase-difference-based synchronization cost. We take the following five steps:

- (i) First, we define the synchronization cost, S_{ij} , due to phase difference between frequency-synchronized phase oscillators i and j in the Kuramoto model. We adopt the Kuramoto model because the model has been used as approximation of various systems including power grids [2, 21, 22] and large-scale brain networks [23, 24].
- (ii) Second, by using the definition, we numerically calculate the mean of synchronization

cost, $\langle S \rangle$, in the entire network.

- (iii) Third, by using a rewiring strategy [25], we show that "rich-club" network topology [24, 26–28], which consists of densely inter-connected modules and peripheral low-degree nodes, is the optimal topology with the minimum synchronization cost.
- (iv) Forth, we characterize the rich-club network topology by quantifying the bimodality of the degree distribution of the network.
- (v) Finally, we provide an analytical interpretation on why the rich-club network organization is associated with a small amount of the synchronization cost.

II. METHOD

A. Definition of synchronization cost

We first define synchronization cost due to phase difference in an unweighted and undirected network which is described by an adjacency matrix A and consists of N phase oscillators. A_{ij} is 1 when oscillators i and j are connected, and A_{ij} is 0 when they are not. According to the Kuramoto model, the phase of the oscillator i , θ_i , is described as

$$\dot{\theta}_i = \omega_i - \epsilon \sum_j A_{ij} \sin(\theta_i - \theta_j), \quad (1)$$

where ω_i is the natural frequency of the oscillator i , and ϵ is a coupling strength. To reduce computational cost for the following rewiring-based optimization, we assume that ϵ is constant for any combination of oscillators. In this Kuramoto model, the synchronization cost between oscillators i and j , S_{ij} is defined based on the phase difference between the connected oscillators as follows:

$$S_{ij} = (\theta_i - \theta_j)^2. \quad (2)$$

Note that the synchronization cost is only defined after the network of the oscillators reaches to a state of frequency synchronization. As described in Appendix A, in power grids, S_{ij} can be regarded as an index that is proportional to power loss due to difference in voltage phase between power plants i and j .

B. Estimation of the mean synchronization cost

Based on the definition of S_{ij} , we numerically estimate S_{ij} for each edge in the following four steps for a given network:

- (i) We set normally-distributed $\{\omega_i\}$ for each node. It is because that previous studies on real networks such as power grids and brain networks have assumed that the natural frequencies of belonging oscillators are symmetrically fluctuating around the averaged frequency [4, 22, 29].
- (ii) Based on the Kuramoto model described in Eq. (1), we numerically estimate frequency-synchronized status, where $\dot{\theta}_i$ becomes a common constant value, Ω , for any i .
- (iii) In the frequency-synchronized status, each oscillator has a different specific phase, θ_i . Based on the set of $\{\theta_i\}$, we then evaluate S_{ij} for each edge.
- (iv) As described in (i), the set of the natural frequency, $\{\omega_i\}$, is fluctuating over time. Thus, the S_{ij} is also fluctuating over time. Therefore, we repeat the procedure (i)-(iii) 200 times with different sets of $\{\omega_i\}$, and obtain 200 different sets of $\{S_{ij}\}$. Then, we average the $\{S_{ij}\}$ over time, obtaining $\langle S_{ij} \rangle$ for each edge. Finally, we average the $\langle S_{ij} \rangle$ across edges, and obtain $\langle S \rangle$ for the entire network.

C. Rewiring-based optimization

To search for the optimal network topology with the least $\langle S \rangle$, we adopt the rewiring method that previous studies used to explore the network topology with the largest synchronizability [25]. We apply the following rewiring-based optimization procedure to a given connected network with N nodes and mean degree of $\langle k \rangle$: At each step, the number of rewired edges is randomly determined based on an exponential distribution. The set of edges to be rewired is also randomly chosen in a given network. After the rewiring, we estimate frequency-synchronized status and obtain $\langle S \rangle_{\text{updated}}$. The attempted rewiring is rejected if the updated network is disconnected. Otherwise, the rewiring is accepted if $\Delta \langle S \rangle = \langle S \rangle_{\text{updated}} - \langle S \rangle_{\text{initial}} < 0$, or with probability $p = \min(1, [1 - (1 - q)\Delta \langle S \rangle / T]^{1/(1-q)})$ where T is a temperature-like parameter and $q = -3$ [25]. The initial rewiring is conducted at $T = \infty$, and, after the first N rewiring, T is set as $(1 - q)(\Delta \langle S \rangle)_{\text{max}}$ where $(\Delta \langle S \rangle)_{\text{max}}$ is the largest $\Delta \langle S \rangle$ in the first N rewiring trials. After

that, T is decreased 10% in every 10 rewiring trials. This estimation process iterated until there is no change in more than 50 successive rewiring steps. We apply this rewiring-based optimization to three different initial networks: an Erdős-Rényi (ER) random model, Watts-Strogatz (WS) model [30], and a Barabási-Albert (BA) model [31] with $N = 50$ and $\langle k \rangle = 4$ [25]. In all the cases, the coupling strength, ϵ , is set to be 0.3. Each set of the natural frequencies of Kuramoto oscillators, $\{\omega_i\}$, is randomly chosen from the normal distribution with an average of 100π and a standard deviation of 1.

During the optimization, we trace the standard order parameter, r , and local synchronizability, r_{local} [32], defined as follows:

$$re^{i\psi} = \frac{1}{N} \sum_j e^{i\theta_j}, \quad (3)$$

$$r_{\text{local}} = \frac{1}{2N_l} \sum_i \sum_{j \in \Gamma_i} \left| \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{t_r}^{t_r + \Delta t} e^{i[\theta_i(t) - \theta_j(t)]} dt \right|, \quad (4)$$

where N_l is the total number of edges, Γ_i is the set of neighbors of node i , and ψ is a global phase. Furthermore, after the optimization is completed, we compare the optimized networks from the different initial networks by estimating the following basic topological properties: mean of shortest path length [33], mean of clustering coefficient [30], mean of betweenness centrality [34], and degree correlation [34]. We conducted ten optimizations of ten different networks for each type of the initial networks, and averaged these basic properties.

- The shortest path length, ℓ_{ij} , is defined as the shortest distance between two nodes i and j [33]. The averaged shortest path length, $\langle \ell \rangle$, is defined as the average value of ℓ_{ij} over all the possible pairs of nodes in the network.
- The clustering coefficient for node i , C_i , measures the local group cohesiveness [30], which is defined as the ratio of the number of links between the neighbors of i and the maximum number of such links. We define the mean clustering coefficient, $\langle C \rangle$, as an average value over all the nodes.
- The betweenness centrality for node i , b_i , is defined as the number of shortest paths between pairs of nodes that pass through a given node [34]. We define the mean betweenness centrality, $\langle b \rangle$, as an average value over all the nodes.
- The degree correlation for a network is defined as the Pearson assortativity coefficient of the degrees, $r_{\text{assortative}}$ [34]. The coefficient enables us to quantify the preference for

high-degree nodes to attach to other high-degree nodes. Networks with this preference show large $r_{\text{assortative}}$.

D. Estimation of rich-club coefficient

We estimate rich-club coefficient, $\Phi(k)$, for both initial networks and optimized networks. According to the previous studies [24, 26–28], the coefficient for each degree k is calculated as

$$\Phi(k) = \frac{2E_{>k}}{N_{>k}(N_{>k} - 1)}, \quad (5)$$

where $E_{>k}$ represents the number of edges among $N_{>k}$ nodes that have more than k degrees. As in previous studies [24, 27, 28], we calculate normalized rich-club coefficients, $\Phi_{\text{norm}}(k)$ through dividing the raw value, $\Phi(k)$, by the mean of rich-club coefficients of 100 random networks (ER models), $\Phi_{\text{random}}(k)$, as follows,

$$\Phi_{\text{norm}}(k) = \frac{\Phi(k)}{\Phi_{\text{random}}(k)}. \quad (6)$$

III. RESULTS

A. Rewiring-based optimization

Fig. 1 shows representative results of the rewiring-based optimizations. $\langle S \rangle$ was decreased from approximately 6.5×10^{-2} to 4.5×10^{-2} even when the initial network structure was different. In all the initial networks, $\langle S \rangle$ reached to a stable status after approximate 400 steps of rewiring. Strikingly speaking, we cannot guarantee that the optimal network was found, but this result suggests that a reasonably robust approximation of the optimal topology was obtained in this method.

During the optimization, the standard order parameter, r , were fluctuating just below 1 during the optimization (a small panel in Fig. 1 A). The local synchronizability, r_{local} , showed the similar fluctuation below 1. Considering the previous studies on these parameters [32], these behaviors of r and r_{local} are considered to be related to the amount of the coupling strength, ϵ . The previous studies [32] have demonstrated that, when the coupling strength is more than 0.2, both of r and r_{local} reach a plateau that is near to 1 regardless of network topology. In the present study, the coupling strength, ϵ , was set at 0.3, because global synchronization is

necessary for the estimation of $\langle S \rangle$. This relatively large coupling strength could result in the saturation of the global and local order parameters, r and r_{local} during the optimization process.

Fig. 1 B shows that the optimized networks for the three different initial networks commonly exhibit a characteristic topology, which has a densely interconnected core nodes and peripheral nodes dangling the core module. The heterogeneous network features were also observed in the basic network properties in the optimized networks (Tab. I). Compared with the initial networks, the optimized networks tended to show larger averaged values of the shortest path length, $\langle \ell \rangle$, betweenness centrality, $\langle b \rangle$, and degree correlation, $r_{\text{assortative}}$. The averaged values of the clustering coefficients, $\langle C \rangle$, were smaller in the optimized networks. These results suggest that, through the optimization process, the network seems to enlarge its heterogeneity.

B. Rich-club organization

This heterogeneous network topology has been reported as "rich-club" organization in a series of previous theoretical and experimental studies [24, 26–28]. The prior literatures have characterized the organization by estimating normalized rich-club coefficients, $\Phi_{\text{norm}}(k)$ described in Eq. 6. If the network has rich-club organization, $\Phi_{\text{norm}}(k)$ should be more than 1, and increase monotonically as k increases.

Fig. 2 shows the comparison in $\Phi_{\text{norm}}(k)$ between initial networks and optimized networks. To clarify the difference in the rich-club coefficients, we adopted as the initial networks larger networks than shown in Fig. 1 B (*i.e.*, $N = 100$, $\langle k \rangle = 4$). Before the optimization, $\Phi_{\text{norm}}(k)$ was not always larger than 1 (*e.g.*, ER and WS models), and did not show monotonic increase along k , which is consistent with a previous study [27]. In contrast, in the optimized networks, the rich-club coefficients were larger than 1 in almost all the range of k , and monotonically increased as k increased. These phenomena were observed commonly among the three different optimized networks that were derived from the three different initial networks. In addition to the appearance of the networks in Fig. 1 B, this estimation of $\Phi_{\text{norm}}(k)$ supports the notion that the networks with rich-club organization has the minimum or a very small amount of synchronization cost due to phase difference among frequency-synchronized oscillators.

C. Bimodal Degree Distribution

As shown in Fig. 1 B, the rich-club network topology consists of high-degree nodes cluster and low-degree peripheral nodes. Therefore, we hypothesized that the topology can be characterized by a bimodal degree distribution. To test the hypothesis, we estimated Wolfson's polarization index, \hat{P} [35]. The Wolfson's index for degree distribution is defined as:

$$\hat{P} = \frac{2 \langle k \rangle}{m} (2 (\langle k \rangle_2 - \langle k \rangle_1) - G), \quad (7)$$

where $\langle k \rangle$ is the mean of the degree, k_i , and m denotes the median of the degree. $\langle k \rangle_1$ and $\langle k \rangle_2$ are the mean values of $\{k_i \mid k_i < m\}$ and of $\{k_i \mid k_i \geq m\}$, respectively. G represents Gini inequality index, which is defined as $G = \frac{1}{2 \langle k \rangle} \sum_{i=1}^N \sum_{j=1}^N |k_i - k_j|$. This Wolfson's polarization index shows the extent of the bimodality of the distribution. If the distribution is completely the same as a uniform distribution, the \hat{P} is 0. If the half of population has nothing and the other half shares everything, the \hat{P} reaches a maximum, 0.25. In the present case, a larger \hat{P} indicates that the network has a more bimodal and bipolarized degree distribution.

We estimated \hat{P} during the rewiring-based optimization. Because \hat{P} can be calculated more accurately for networks with more nodes, we used the BA model with $N = 100$ and $\langle k \rangle = 4$ as an initial network for the optimization. As a result, in the course of the above-mentioned optimization, $\langle S \rangle$ decreased during the rewiring-based optimization (Fig. 3A). Meanwhile, as $\langle S \rangle$ decreases, \hat{P} increases (Fig. 3B). Actually, the degree distribution changed from a power-law distribution (Fig. 3C) to a bimodal distribution (Fig. 3D). This relation was also observed for different initial networks (e.g., ER model). This correlation supports the hypothesis that rich-club network with small $\langle S \rangle$ can be characterized by its bimodal degree distribution.

D. Analytical Interpretation

We finally provide an analytical interpretation on why the rich-club network has less $\langle S \rangle$. Using mean-field approximation, the Eq. (1) in the frequency-synchronized status can be described as $\Omega = \omega_i - \epsilon k_i \sin(\theta_i - \psi)$, where ψ is defined in the Eq. (3). Therefore, if $|\theta_i - \psi|$ is small enough, $\theta_i - \psi = \frac{1}{\epsilon k_i}(\omega_i - \Omega)$, and $(\theta_i - \theta_j)^2$ is described as

$$(\theta_i - \theta_j)^2 = \frac{(\omega_i - \Omega)^2}{\epsilon^2 k_i^2} + \frac{(\omega_j - \Omega)^2}{\epsilon^2 k_j^2} - \frac{2(\omega_i - \Omega)(\omega_j - \Omega)}{\epsilon^2 k_i k_j}, \quad (8)$$

for a set of $\{\omega_i\}$. $\langle S_{ij} \rangle$ is obtained as averaged $(\theta_i - \theta_j)^2$ across an enough number of sets of $\{\omega_i\}$. As in the above-described numerical estimation, we assume that $\{\omega_i\}$ is distributed according to the normal distribution with a mean of ω_0 and a standard deviation of σ , and that the synchronized frequency is always Ω in every set of $\{\omega_i\}$. Because we can also assume that ω_0 is nearly equal to Ω , $\langle (\omega_i - \Omega)^2 \rangle$ is considered to be equal to σ^2 , and $\langle (\omega_i - \Omega)(\omega_j - \Omega) \rangle$ is considered to be equal to zero. Consequently, we obtain the approximation of $\langle S_{ij} \rangle$ as follows:

$$\langle \tilde{S}_{ij} \rangle = \frac{\sigma^2}{\epsilon^2} \left(\frac{1}{k_i^2} + \frac{1}{k_j^2} \right). \quad (9)$$

This approximation was validated through comparison of $\langle \tilde{S}_{ij} \rangle$ with the real $\langle S_{ij} \rangle$, as shown in Fig. 4 A ($\sigma = 1, \epsilon = 0.3$). The two parameters had a large negative value of Pearson's correlation coefficient (-0.88).

This expression of $\langle \tilde{S}_{ij} \rangle$ gives qualitative explanation on relationship between rich-club network topology and a small value of $\langle S \rangle$: To achieve a small amount of $\langle S \rangle$, $\langle \tilde{S}_{ij} \rangle$ should be small. When the node i has a high degree, it makes more contribution to a smaller $\langle \tilde{S}_{ij} \rangle$ to connect with a the node j with a high degree. It is also case when the node i has a small degree. As a result, for a small value of $\langle \tilde{S}_{ij} \rangle$, high degree nodes should connect with other high degree nodes, and low degree nodes should not have edges with other low degree nodes, but with high degree nodes. Consequently, high degree nodes tend to be gathered and create a densely-connected core module, and low degree nodes tend to connect with high degree nodes in the core module. Overall, the optimized network with a small amount of $\langle S \rangle$ is likely to be a rich-club network.

Note that it is difficult to further extend this approximation. If this approximation of $\langle S_{ij} \rangle$ is accurate enough, a simple calculation of Eq. 9 leads us to the proportional relationship between $\langle S \rangle$ and $\frac{1}{\langle k \rangle} \langle \frac{1}{k} \rangle$. Given $\langle k \rangle$ is a constant value as in the present study, $\langle S \rangle$ should be proportional to $\langle \frac{1}{k} \rangle$. However, as shown in Fig. 4 B, we could not observe a linear relationship. This inaccurate approximation of $\langle S \rangle$ may be caused by accumulation of the small difference between $\langle S_{ij} \rangle$ and $\langle \tilde{S}_{ij} \rangle$. This result suggests that we cannot extend this approximation to represent $\langle S \rangle$ only by $\langle \frac{1}{k} \rangle$.

IV. DISCUSSION

The present study introduced synchronization cost based on phase difference among frequency-synchronized oscillators. Using the rewiring-based optimization [25, 36], we showed that the synchronization cost is minimized in a rich-club network topology. Furthermore, we demonstrated that the network topology can be characterized by the bimodality of its degree distribution. Finally, we provided analytical explanation on why the rich-club network topology is associated with a small amount of synchronization cost.

The concept of synchronization cost is not a novel idea of the present study. As described in Sec. I, a line of previous studies have investigated a different type of synchronization cost, which is based on coupling strength [17–19]. Whereas the present synchronization cost due to phase difference can be regarded as dynamic cost per unit time, the cost based on coupling strength can be considered as static cost that is related to building and maintaining of network infrastructures. Interestingly, the optimal network topology with the least cost depends on which of the two types of synchronization cost we adopt. The optimal networks for the synchronization cost based on coupling strength often show more homogeneous properties [17] than those for the other synchronization cost. The homogeneity of networks is desired to enhance synchronizability [25, 37]. Therefore, it may be necessary to investigate what network structures balance these two types of synchronization cost.

The present synchronization cost in the present study can be another concept of load assigned to edge in a complex network. Previous studies used edge-betweenness as edge load [38, 39], which is useful in various situations from human interaction [39] to data transmission in computer networks [38]. However, because the edge betweenness does not consider synchrony in networks, its properties have evident difference from those of the synchronization cost. For example, as shown in Sec. IIID, the synchronization cost is lower between high degree nodes, and higher between low degree nodes. In contrast, the edge betweenness tends to be higher in edges bridging high degree nodes and be lower in edges bridging low degree nodes. These distinct properties suggest the possibility that the synchronization cost can be another concept of edge load.

The synchronization cost in the present study, S_{ij} , has a mathematical expression similar to that for local synchronizability, r_{local} [32]. However, the two parameters focus on different phases of synchronization in complex networks. The local synchronizability enables us to quantify

the local construction of the synchronization pattern. Therefore, it is useful to investigate properties of networks that are not yet fully synchronized. In contrast, the synchronization cost in the present study can be only estimated in fully-synchronized networks. Therefore, in the present study, we used a relatively large coupling strength, and achieved full synchronization throughout the optimizations. As a result, in the present study, the local synchronizability was always saturated.

Although the present study did not adopt models specific to any real networks, the findings may help understanding large-scale brain networks. Recently, a few of recent studies have reported the existence of the rich-club organization in the large-scale brain networks. A previous empirical study has demonstrated the existence of the rich-club organization in the large-scale human brain networks [28]. Another study has investigated the anatomical connectivity in the cerebral cortex of cats, and has showed that rich-club organization controls the dynamic transition of synchronization in the brain [24]. A recent review has suggested that the organization is a cost-effective network topology for the brain networks, which are required to be adapted to various cognitive functions [40]. In addition to the context of cost-effectiveness, the rich-club network topology is robust to random attack [41]. This previous study analytically and numerically demonstrated that networks robust to random attack have similar structures observed in the present study. The robust networks highly-interconnected hub modules and peripheral nodes (*leaf* nodes) that have a single edge. This network topology has bimodal degree distribution and shows rich-club organization. Considering these prior literatures, it is suggested that the rich-club organization is beneficial for the large-scale brain networks to efficiently and robustly maintain its wide range of functions based on synchronization.

Acknowledgments

The author acknowledges the support from the Japan Society for the Promotion of Science (JSPS) Research Fellowship for Young Scientists (222882).

-
- [1] O. I. Elgerd, *Electric energy systems theory* (McGraw-Hill Companies, 1982); P. S. Kundur, *Power system stability and control* (McGraw-Hill Companies, 1993).
 - [2] F. Dörfler and F. Bullo, American Control Conference 2010, 930-937 (2010).

- [3] F. Varela, J.P. Lachaux, E. Rodriguez, and J. Martinerie, Nat Rev Neurosci. **2(4)**:229-39 (2001).
- [4] J. Fell, and N. Axmacher, Nat Rev Neurosci. **12(2)**:105-18 (2011).
- [5] D.J. Hill, and G. Chen, IEEE Int. Symposium on Circuits and Systems 711-715 (2006).
- [6] H. Gharavi, and R. Ghafurian, Proceedings of the IEEE **99**, 917-921 (2011); G.W. Arnold, *ibid.* **99**, 922-927 (2011).
- [7] Energy Future Coalition, Report of Smart Grid Working Group, (2008).
- [8] S.L. Bressler, R. Coppola, and R. Nakamura, Nature. **366(6451)**:153-6 (1993).
- [9] P.R. Roelfsema, A.K. Engel, P. König, and W. Singer, Nature. **385(6612)**:157-61 (1997).
- [10] W.H. Miltner, C. Braun, M. Arnold, H. Witte, E. Taub. Nature. **397(6718)**:434-6 (1999).
- [11] J.F. Hipp, A.K. Engel, and M. Siegel, Neuron. **69(2)**:387-96, (2011).
- [12] R.D. Traub, M.A. Whittington, I.M. Stanford, and J.G. Jefferys, Nature. **383(6601)**:621-4 (1996).
- [13] T. Womelsdorf, J-M. Schoffelen, R. Oostenveld, W. Singer, R. Desimone, A.K. Engel, and P. Fries, Science. **316(5831)**:1609-12 (2007).
- [14] P. Fries, Trends Cogn Sci (Regul Ed). **9(10)**:474-80 (2005).
- [15] M. Stopfer, S. Bhagavan, B.H. Smith, and G. Laurent, Nature. **390(6655)**:70-4 (1997).
- [16] G. Tononi, and G.M. Edelman, Brain Res Brain Res Rev. **31(2-3)**:391-400 (2000).
- [17] A.E. Motter, C. Zhou, and J. Kurths, Phys. Rev. E **71**, 016116 (2005).
- [18] T. Nishikawa, and A.E. Motter, Physica D **224**, 77-89 (2006).
- [19] M. Brede, Eur. Phys. J. B. **62**, 87-94 (2008).
- [20] A. Arenas, A. Diaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Phys. Rep. **469**, 93-153 (2008).
- [21] L. Buzna, S. Lozano, and A. Diaz-Guilera, Phys. Rev. E **80**, 066120 (2009).
- [22] V. Fioriti, S. Ruzzante, E. Castorini, E. Marchei, and V. Rosato, Critical Information Infrastructure Security 14-23 (2009).
- [23] M. Breakspear, S. Heitmann, and A. Daffertshofer, Front. Hum. Neurosci. **4**:190 (2010).
- [24] J. Gómez-Gardeñes, G. Zamora-López, Y. Moreno, and A. Arenas. PLoS ONE. **5(8)**:e12313 (2010).
- [25] L. Donetti, P.I. Hurtado, and M.A. Muñoz, Phys. Rev. Lett. **95**, 188701 (2005).
- [26] S. Zhou, and R. Mondragon. IEEE Commun Lett. **8(3)**:180-2 (2004).
- [27] V. Colizza, A. Flammini, M. Serrano, and A. Vespignani, Nat Phys. **2(2)**:110-5 (2006).
- [28] M.P. van den Heuvel, and O. Sporns, Journal of Neuroscience. **31(44)**:15775-86 (2011).

- [29] G. Filatrella, A.H. Nielsen, and N.F. Pedersen, Eur. Phys. J. B. **61**, 485-491 (2008).
- [30] D.J. Watts and S.H. Strogatz, Nature **393**, 440 (1998).
- [31] A.L. Barabási and R. Albert, Science **286**, 5439 (1999).
- [32] J. Gómez-Gardeñes, Y. Moreno, and A. Arenas, Phys. Rev. Lett. **98** 034101 (2007), Phys. Rev. E. **75** 066106 (2007).
- [33] A. Bunde, and S. Havlin, *Fractals and Disordered System* (Springer Verlag, 1991).
- [34] M.E.J. Newman, Phys. Rev. E. **64** 016131 (2001).
- [35] M. Wolfson, Review of Income and Wealth **43**, 4 401 (1997).
- [36] TE Gorochoowski, M di Bernardo, CS Grierson. Phys Rev E. **81** 056212 (2010).
- [37] T. Nishikawa, A.E. Motter, Y.C. Lai, and F.C. Hoppensteadt, Phys. Rev. Lett. **91**, 014101 (2003).
- [38] P. Holme, Phys. Rev. E **66**, 036119 (2002).
- [39] R. Guimerà, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas, Phys. Rev. E **68**, 065103 (2003).
- [40] E. Bullmore, and O. Sporns, Nat Rev Neurosci. 336-349, (2012).
- [41] G. Paul, S. Sreenivasan, S. Havlin, and H.E. Stanley, Physica A. **370(2)**:854-62 (2006).

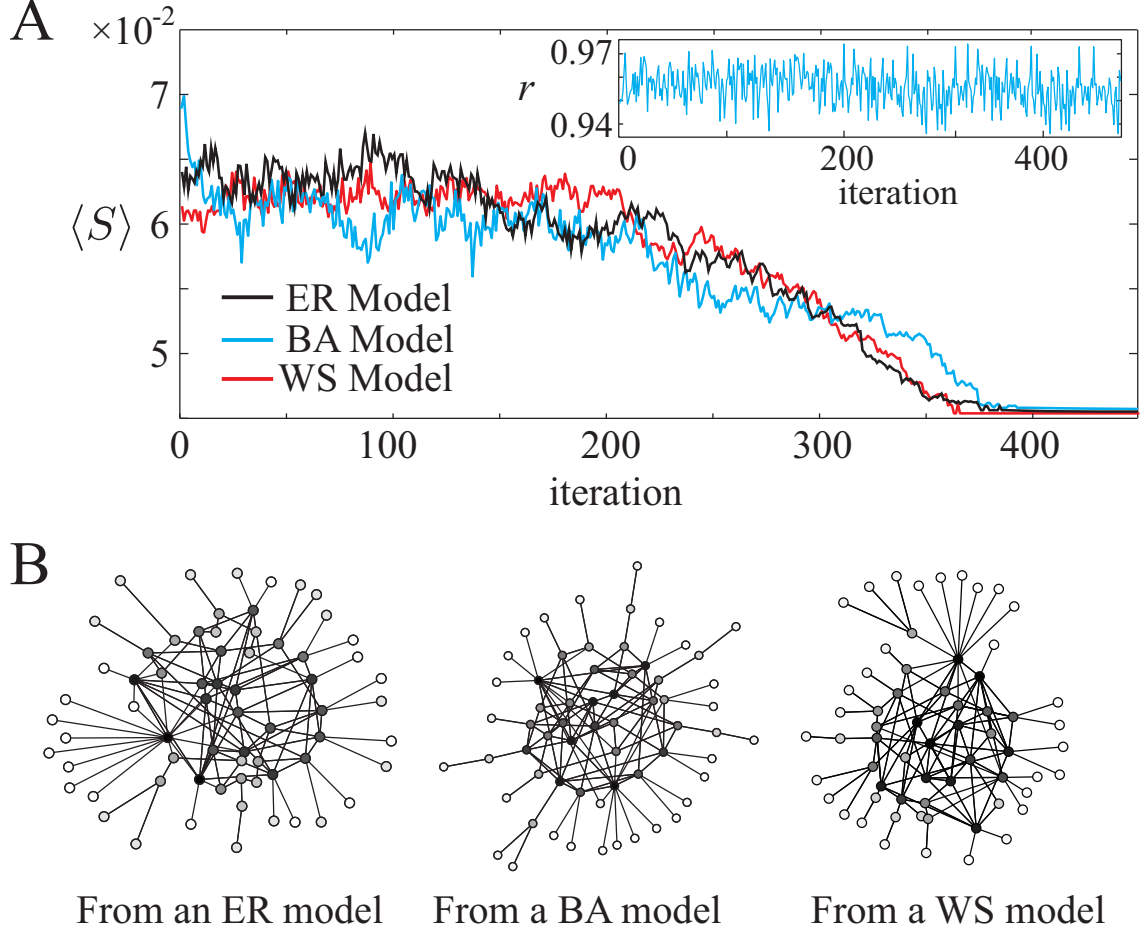


FIG. 1: (Color online) **A**. Main panel: Change of the synchronization cost, $\langle S \rangle$, during rewiring-based optimization. Despite different initial networks (ER, BA, and WS models) with $N = 50$, $\langle k \rangle = 4$, the synchronization cost converge to a similar amount of $\langle S \rangle$. Sub panel: Change of the standard order parameter, r , during the optimization. In contrast to $\langle S \rangle$, the standard order parameter does not show notable change, just fluctuating below 1. The line shows the change of r when the initial network is the BA model. In cases of the other two initial networks, the similar fluctuations were observed. **B**. Network topology optimized from different initial networks. Optimized networks are similar to each other. They have rich-club network topology, which consists of a densely-connected core module and peripheral low degree nodes connecting with the core. The color in the nodes represent the degree of the nodes: darker nodes have more edges.

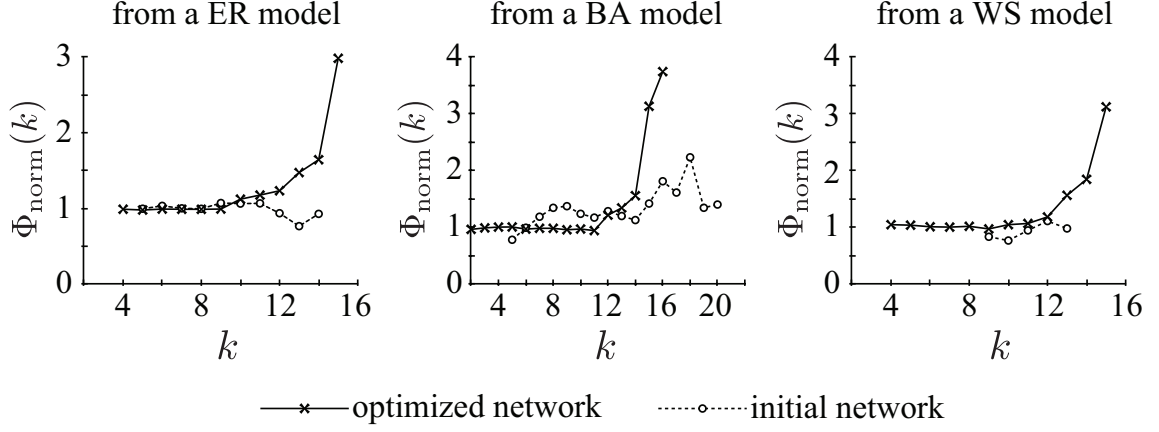


FIG. 2: Difference in rich-club coefficients between the initial networks (circles and dashed lines) and optimized networks (multiple marks and solid lines). While the normalized rich-club coefficients $\Phi_{\text{norm}}(k)$ do not show monotonic increase in the initial networks, those in the optimized networks monotonically increase. These results suggest that the rewiring-based optimization changes the initial networks to networks with rich-club organization. To clarify the difference between before and after optimization, we adopted larger networks ($N = 100, \langle k \rangle = 4$) than in Fig. 1 ($N = 50, \langle k \rangle = 4$).

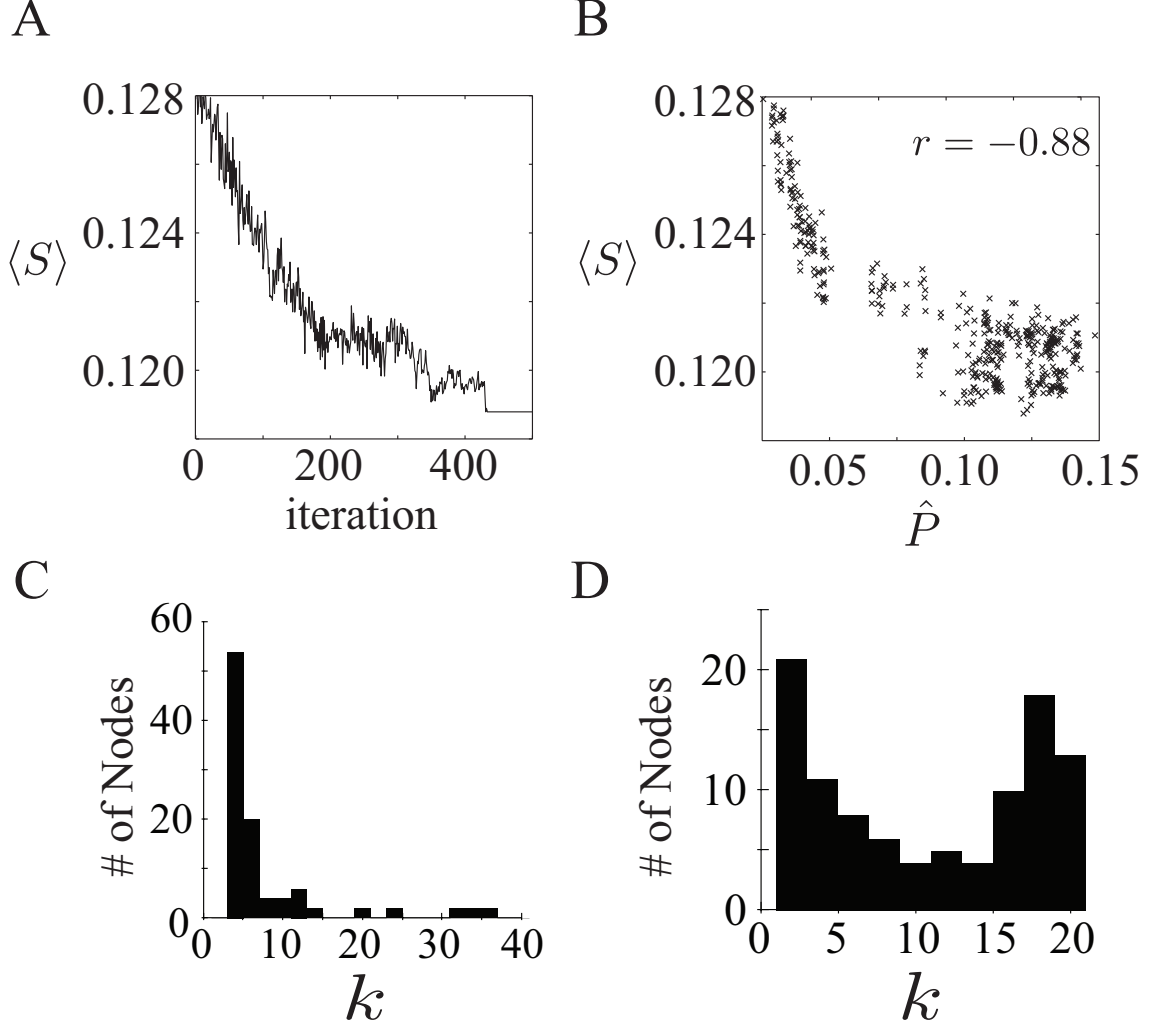


FIG. 3: As the synchronization cost, $\langle S \rangle$, decreases in the rewiring-based optimization (panel **A**), the Wolfson's polarization index, \hat{P} , increases (panel **B**). This relation suggests that the rich-club network with a small amount of the synchronization cost can be characterized by bimodal degree distribution, which is quantified by Wolfson's polarization index. Indeed, the degree distribution changed from a power-law distribution (panel **C**) to a bimodal distribution (panel **D**). To clarify the difference between before and after optimization, we adopted larger networks ($N = 100, \langle k \rangle = 4$) than in Fig. 1 ($N = 50, \langle k \rangle = 4$).

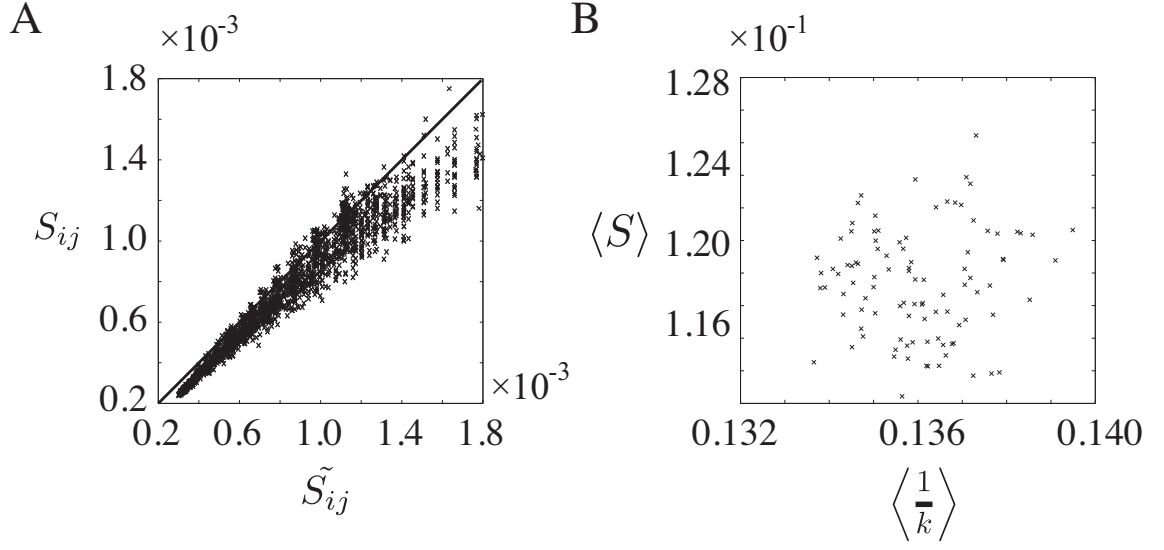


FIG. 4: **A** The analytical approximation of synchronization cost, $\langle \tilde{S}_{ij} \rangle$, is predictive of the real synchronization cost, $\langle S_{ij} \rangle$, in the BA model with $N = 200, \langle N \rangle = 10$. **B** Simple calculations using $\langle \tilde{S}_{ij} \rangle$ suggest a positive linear relationship between $\langle S \rangle$ and $\langle \frac{1}{k} \rangle$. However, there is not a strong correlation between them, which suggests a limitation of the approximation.

TABLE I: Basic topological properties of the optimized networks. Despite different initial networks, the optimized networks had similar network topological properties. The values for the initial networks represent averaged values across ten estimations, whereas the values for the optimized networks show the mean \pm s.d. across the ten estimations.

	$\langle \ell \rangle$		$\langle C \rangle$		$\langle b \rangle$		$r_{\text{assortative}}$	
	initial	optimized	initial	optimized	initial	optimized	initial	optimized
From ER models	1.9	3.5 ± 0.025	0.21	0.13 ± 0.010	44	122.2 ± 3.7	0.034	0.25 ± 0.12
From BA models	1.8	3.3 ± 0.021	0.33	0.15 ± 0.012	48	113 ± 2.8	-0.15	0.28 ± 0.021
From WS models	2.3	3.1 ± 0.017	0.62	0.16 ± 0.010	63	105 ± 2.4	0.047	0.25 ± 0.011

Appendix A: Definition of Synchronization Cost

In this section, we explain why power loss consumed in the electric line between power plants can be represented by square of difference in phase of voltage between the two power plants.

In the following model, as in previous studies [2, 22], we do not consider the effect of the length of the power line on the power loss. To estimate the power loss in a typical power line shown in Fig. 5, we estimate active power flow (P_{ij} and P_{ji}), reactive power flow (Q_{ij} and Q_{ji}), and delayed reactive power flow (Q_{ci} and Q_{cj}) as follows [1]:

$$P_{ij} = \frac{V_i V_j \sin(\theta_i - \theta_j)}{Z_{ij}^2 / f_0 L_{ij}} + \frac{V_i^2 - V_i V_j \cos(\theta_i - \theta_j)}{Z_{ij}^2 / R_{ij}}, \quad (\text{A1})$$

$$Q_{ij} = -\frac{V_i V_j \sin(\theta_i - \theta_j)}{Z_{ij}^2 / R_{ij}} + \frac{V_i^2 - V_i V_j \cos(\theta_i - \theta_j)}{Z_{ij}^2 / f_0 L_{ij}}, \quad (\text{A2})$$

$$Q_{ci} = \frac{f_0 C_{ij}}{2} V_i^2, \quad (\text{A3})$$

where f_0 represents synchronized angular frequency of alternating voltage, and $Z_{ij}^2 = R_{ij}^2 + (f_0 L_{ij})^2$. P_{ji} , Q_{ji} , and Q_{cj} are obtained by exchanging i and j . Using these power flows, the active power loss due to resistance, P_{loss}^{ij} , is calculated as $P_{ij} + P_{ji}$, whereas the reactive power loss due to inductance, Q_{loss}^{ij} , is estimated as $Q_{ij} + Q_{ji} + Q_{ci} + Q_{cj}$ as follows:

$$P_{\text{loss}}^{ij} = \frac{R_{ij}}{Z_{ij}^2} (-2V_i V_j \cos(\theta_i - \theta_j) + V_i^2 + V_j^2), \quad (\text{A4})$$

$$Q_{\text{loss}}^{ij} = \frac{Z_{ij}^2}{f_0 L_{ij}} (-2V_i V_j \cos(\theta_i - \theta_j) + V_i^2 + V_j^2) + \frac{f_0 C_{ij}}{2} (V_i^2 + V_j^2). \quad (\text{A5})$$

The total power loss is estimated as a combination of the active power loss and the reactive power loss [1]. By using a second-order Taylor expansion, we regard the total power loss, $P_{\text{loss}}^{ij} + Q_{\text{loss}}^{ij}$, as $a_0 + a_1 (\theta_i - \theta_j)^2$, where a_0 and a_1 are constants ($a_1 > 0$). Therefore, we define the synchronization cost, S_{ij} , for a power line between power plants i and j as

$$S_{ij} = (\theta_i - \theta_j)^2. \quad (\text{A6})$$

Appendix B: Power Grid as Kuramoto Model

In this section, we explain that, under several assumptions, we can approximate power grids by the first-order Kuramoto model of nonuniform oscillators.

As previous studies [2, 21, 22], we model a power grid as follows: The structure of the power grid with N power plants is represented as an unweighted and undirected adjacency matrix A ,

where a node represents a power plant and an edge a power line. A_{ij} is 1 when power plants i and j have a power line between them, and A_{ij} is 0 when they do not. According to the previous studies [2, 21, 22], the phase of the output voltage of the power plant i , θ_i , is described as

$$\dot{\theta}_i = \frac{f_i}{D_i} - \sum_j \frac{W_{ij}}{D_i} A_{ij} \sin(\theta_i - \theta_j), \quad (\text{B1})$$

where D_i denotes a damping constant, W_{ij} is an amount of power transfer between power plants i and j , and f_i represents the natural frequency of the output voltage from the power plant i . To reduce computational cost for the following rewiring-based optimization, we assume that $W_{ij}/D_i = \epsilon$ for any power line. Because f_i/D_i is specific to power plant i , we replace the value with ω_i . Consequently, the voltage phase of the power plants can be expressed in the Kuramoto model as

$$\dot{\theta}_i = \omega_i - \epsilon \sum_j A_{ij} \sin(\theta_i - \theta_j). \quad (\text{B2})$$

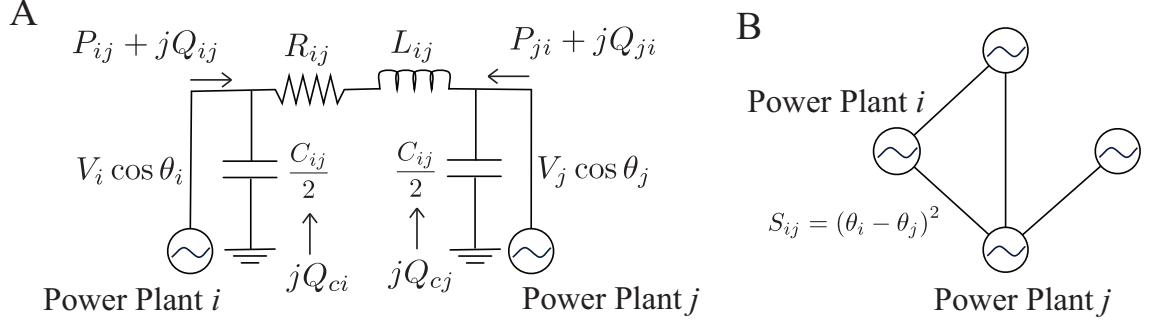


FIG. 5: Panel **A** shows a typical power line between power plants *i* and *j*. $V_i \cos \theta_i$ and $V_j \cos \theta_j$ indicate the voltage of the output from the power plants. R_{ij} , L_{ij} , and C_{ij} indicate resistance, inductance, and conductance between the power plants. As show in panel **B**, S is defined in every power line based on the phase difference of the voltages between the connecting power plants.